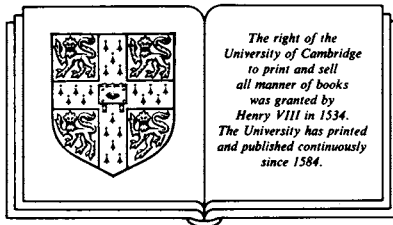

Physical analysis for tribology

Terence F. J. Quinn

Professor of Engineering
School of Engineering and Applied Science
United States International University
(European Campus)
Bushey, Hertfordshire



CAMBRIDGE UNIVERSITY PRESS

Cambridge

New York Port Chester

Melbourne Sydney

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
40 West 20th Street, New York, NY 10011, USA
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

© Cambridge University Press 1991

First published 1991

British Library cataloguing in publication data

Quinn, Terence F. J.
Physical analysis for tribology.
1. Tribology
I. Title
621.8'9

Library of Congress cataloguing in publication data available

ISBN 0 521 32602 8 hardback

Transferred to digital printing 2003

Contents

	<i>page</i>
Preface	vii
Nomenclature list	ix
1 Tribology	1
1.1 Definitions of common tribological terms	1
1.2 Frictional behaviour of materials	9
1.3 Lubrication between surfaces in relative motion	19
1.4 The wear of sliding and rolling surfaces	22
1.5 Concluding remarks on the relevance of physical analytical techniques for tribology	34
2 Macroscopic physical techniques	35
2.1 Optical microscopy and optical interferometry	35
2.2 Transmission electron microscopy of replicas of surfaces	50
2.3 Profilometry and scanning electron microscopy (SEM)	56
2.4 Electrical capacitance and resistance	64
2.5 Magnetic perturbation techniques	67
2.6 Macroscopic techniques for monitoring oil-lubricated machinery (with special emphasis on 'ferrography')	77
2.7 Concluding remarks regarding the importance of macroscopic physical techniques to tribology	80
3 Microscopic physical techniques for studying atomic arrangement	82
3.1 Basic crystallography	82
3.2 X-ray diffraction techniques	98
3.3 Electron diffraction methods	123
3.4 Concluding remarks regarding the importance of microscopic techniques for studying the atomic arrangements in tribo-elements	146
4 Microscopic techniques for studying atomic structure	148
4.1 Optical spectra	148
4.2 X-ray spectra	153
4.3 Electron probe microanalysis	166
4.4 X-ray photoelectron spectroscopy	175
4.5 Auger electron spectroscopy	190
4.6 Other microscopic techniques	200
4.7 Concluding remarks	201

5	The analysis of lubricant films	202
5.1	The analysis of extreme-pressure lubricant films formed in the presence of typical disulphide additives	202
5.2	The analysis of solid lubricant films formed when lamellar solids slide against metals and alloys	224
5.3	The application of physical methods of analysis to the measurement of lubricant film thickness	259
5.4	The analysis of surface films formed in the presence of anti-wear additives	275
6	The analysis of surface temperatures in tribo-systems	291
6.1	The thermal aspects of sliding	291
6.2	The crystallographic changes caused by 'hot-spot' temperatures during sliding wear	313
6.3	Direct measurement of surface temperatures	338
6.4	Concluding remarks on the relevance of physical analytical techniques for analysing surface temperatures in tribo-systems	358
7	The analysis of pitting failures in tribo-systems	360
7.1	Failure in rolling contact bearings in nominal point contact	360
7.2	Basic research into the contact fatigue of ball-bearings	363
7.3	Failures in gears and other tribo-systems in nominal line contact	365
7.4	Basic research on the pitting of discs	368
7.5	Physical analysis of contact fatigue pitting in tribo-systems	376
7.6	The current position in contact fatigue pitting failures	397
8	The analysis of oxidational wear in tribo-systems	399
8.1	Introduction	399
8.2	The oxidational wear of a typical low-alloy steel tribo-system	400
8.3	The oxidational wear of the high-chromium ferritic steel/austenitic stainless steel tribo-system	408
8.4	The oxidational wear of 316 stainless steel tribo-systems	417
8.5	Concluding remarks	431
9	The application of physical techniques to selected ceramic tribo-systems	432
9.1	Introduction	432
9.2	Silicon nitride tribo-systems	432
9.3	Silicon carbide tribo-systems	439
9.4	Sapphire tribo-systems	446
9.5	Concluding remarks on the application of physical techniques to ceramic tribo-systems	456
	References	458
	Index	477

I Tribology

I.1 Definitions of common tribological terms

I.1.1 Tribology

Tribology is a new word based on the Greek word 'tribo', which means 'rubbing'. Hence tribology is the 'study of rubbing'. The word was first used by a British Government committee (chaired by Dr Peter Jost and hence known as the 'Jost Committee') that produced a report, in 1966, calling for increased education and research into a subject that was estimated (at 1966 prices) to be costing the United Kingdom about £300 million per year.

The Jost Committee defined tribology as 'the study of the science and technology of interacting surfaces in relative motion'. It was hoped that the new word might provide the basis of a more unified approach to subjects previously studied separately under titles such as 'friction', 'adhesion', 'lubrication' and 'wear'. It is indeed unfortunate that, to date, the Jost Committee's awareness of the need for a more unified (that is, interdisciplinary) approach has not been shared by many tribologists. This resistance to the calls for a change in our partisan approaches to the subject is illustrated quite neatly by the fact that it has taken nearly 20 years for the *Journal of Lubrication Technology (JOLT)* to change its name to the *Journal of Tribology (JOT)*, namely from 1966 to 1985. It is hoped that this change in name will open the journal to a new generation of papers written by physicists, chemists, chemical engineers and materials engineers (each with their own particular approach to tribology), as well as mechanical engineers (with their special interests in rheology and lubricant pressures).

I.1.2 Friction

The earliest studies of relative motion between two contacting bodies were carried out by Leonardo da Vinci, in the fifteenth century. He showed that the tangential force needed to instigate sliding between two loaded surfaces was proportional to the normal load being applied at the contact, but was independent of the apparent area of that contact. These are the laws of friction, where friction is the resistance to shearing force between the two solid bodies. The ratio of the frictional force to the normal force is called the 'coefficient of friction'. For surfaces sliding without lubrication, this ratio tends to take on values between about 0.3 and 0.9. Leonardo was, in fact, measuring the static friction, whereas most modern tribologists are more interested in the tangential force required to maintain motion, that is the dynamic friction. For surfaces that roll with very little or no sliding, the frictional force is much less. This type of friction is known as 'traction' and is normally measured under lubricated conditions.

1.1.3 Lubrication

The study of lubrication had reached a fairly advanced stage at the time the new word 'tribology' was introduced in the 1960s. Consequently, much of the definitive work in tribology has occurred in lubrication; in particular, the lubrication of gears, cams and tappets and other non-conforming contacts. The prime task of a lubricant is to prevent damage of the interacting surfaces, while they are performing their designed function. For gears, this function is the transmission of power. For plain bearings, the function is to support the load or, sometimes, to provide location, by restricting the number of degrees of freedom. A secondary, but equally important, task of the lubricant is to provide a means of cooling the interacting surfaces. A lubricant delivery system is often required as an aid to this cooling, which thereby provides a means of removing extraneous particles in the lubricant by means of suitable filtration.

Lubrication theory has tended to concentrate on the complexities involved in applying corrections to the classical equation of hydrodynamics, namely Reynold's equation. These corrections arise from the very strong dependence of the coefficient of viscosity upon the *pressures* generated in the non-conforming contact, such as between gear teeth and between cams and tappets. There has been considerable progress in formulating the elastic and hydrodynamic equations governing these pressures as a function of position within the contact zone. This is the topic known as 'elasto-hydrodynamics'. There is, however, an equally strong influence of temperatures within the contact zone upon the coefficient of viscosity of the lubricant. Very little *new* research has been carried out on this aspect of the lubrication of non-conforming geometries.

1.1.4 Wear

Neale (1973) has defined wear as the progressive loss of substance resulting from *mechanical* interaction between two contacting surfaces. In general, these surfaces will be in relative motion, either sliding or rolling, and under load. This definition is too restricted. It should include the possibility that wear can occur through the combined effects of *chemical reaction* with the ambient fluid (whether that fluid be air or oil) *and* mechanical interaction. It should also distinguish between the deliberate removal of one surface (e.g. the workpiece) and the unwanted wear of the other surface (i.e. the cutting tool). Even for two specimens (*A* and *B*) of *identical bulk* composition and hardness, where one might expect equal wear on either face, it is found that differences in the rate of surface removal may occur, especially where there exist differences in *nominal* surface configurations. Different surface configurations will lead to different surface temperatures (T_{SA} and T_{SB}) which, in turn, lead to (a) differences in oxidation states of both surfaces (if metallic); (b) differences in surface and substrate hardnesses (if metallic or polymeric); and (c) differences in the type and amount of interaction between the lubricant and the oxide film on the two surfaces (if metallic or ceramic).

From the previous paragraph, the reader may have realized that wear is really only definable in terms of specimen geometry, specimen hardness and a precise knowledge of the ambient conditions under which the tribo-system is wearing.

There does not exist a single number which *universally* holds valid as the number of cubic metres removed per metre sliding distance for a given pair of surfaces. When a wear rate is quoted, it should also be made clear to which surface (*A* or *B*) this rate is relevant. The surface temperature (T_{SA} or T_{SB}) should also be quoted, together with the hardness value concomitant with this temperature. The constitution of the atmosphere or fluid around each surface should also be quoted in terms of its oxygen and additive content.

1.1.5 Pitting

For many years the materials engineer has been aware that repeatedly compressing and stretching a metal in a typical rotating-bending machine will cause rupture and fatigue, even though the stresses are well below the yield strength. A similar alternation of elastic stretching and compressing occurs in the contact of gear teeth and cams and tappets. Essentially, there is always a position in the relative movement between the opposing members where *pure* rolling occurs. On either side of that extremely narrow region of pure rolling, relative sliding will occur. It is the combination of sliding contact and rolling contact that can sometimes cause a failure of one of the surfaces by pitting, that is a characteristic cratering of the surface, due to the removal of relatively large pieces of that surface by a contact fatigue process. As the surfaces deform each other elastically (in the rolling contact region) over many millions of cycles (for the tribo-system using oil as the lubricant) or many thousands of cycles (for tribo-systems lubricated with water emulsions) fatigue cracks are initiated, which eventually lead to fracture and the formation of pits. Much work has been carried out on contact mechanics (the analysis of the contact stresses occurring in tribo-systems), and even more on the fatigue of materials, but very little has been done towards bringing these two relevant areas together to describe 'contact fatigue' analytically. Pitting, or contact fatigue, is still awaiting its complete elucidation.

1.1.6 Fretting wear

This is sometimes known as 'fretting corrosion', mainly because this type of wear manifests itself through the formation of oxides and/or hydroxides in the region of the fretting contact. It was originally thought that this implied that corrosion had occurred rather than wear. Essentially, fretting wear arises due to the *microscopic* oscillations that occur between two surfaces in contact (whether that contact merely be due to the weight of one of the members of the system or a nut and bolt) when the whole system is subject to vibration. For amplitudes of vibration greater than about $10\mu\text{m}$, it is usual to refer to the system as one in which 'reciprocated sliding' is taking place. Possibly the most important characteristic of fretting is the fact that the wear debris is trapped between the vibrating surfaces and hence plays an important part in the subsequent wear processes. Fretting wear is normally only a nuisance as, for instance, in the fretting damage caused during the transportation of cars on transporters from the factory to the dealer. Even with the engine and wheel bearings fully lubricated, it is not normally possible to exclude all metal-to-metal contact, especially after

being subjected to several hours of vibration on the back of the transporting vehicle.

One of the most irritating things about fretting is its unexpected occurrence in situations not normally considered to be tribosystems. In the Advanced Gas-Cooled Nuclear Reactor (the AGR), for instance, the clamps that hold the cooling rods can suffer fretting damage arising directly from the vibrations caused by the passage of the cooling gas (a mixture of CO and CO₂). This damage reduced the replacement time from the projected 30 years down to 2 or 3 years.

1.1.7 Abrasion

'Abrasion', as normally perceived by the engineer on the macroscopic scale, truly is 'the removal of substance resulting from mechanical interaction between two contacting surfaces' (Neale, 1973). This definition of wear describes 'abrasion' most adequately, especially where it involves 'two body abrasive' contact. This is where one of the bodies is so hard (e.g. a diamond wheel) that all the material removed originates from the workpiece. Sometimes, a third body becomes involved either deliberately, for an abrasive grit embedded into a lead lap, or accidentally, in the case of a very hard piece of one of the surfaces (perhaps an oxide) acting analogously to the abrasive grit particle. 'Three body abrasion' is considered to be rather complex and is, therefore, not the subject of very much reliable hypothesising. Abrasion of drills and mechanical diggers by rock and soil and abrasion of the guideways and belts of mineral conveyor systems by the minerals they are conveying are just a few examples of the relevance of abrasion to industry, particularly the construction industry. Because of their size, the economic losses incurred by abrasive wear of systems like the mechanical digger form a significant part of the losses due to all types of wear. Typically, however, these systems do not readily lend themselves to simulation in the laboratory, nor to the application of physical analytical techniques. In fact, it will be shown (in Section 1.4) that abrasion is a wear *mechanism* rather than a wear *classification*, a mechanism that seems to be involved in some form or another in several types of wear. We will see that physical techniques can be effectively used in elucidating these abrasive wear mechanisms when they occur on the microscopic scale.

1.1.8 Adhesion

Strictly speaking, adhesion is a term relating to the force required to separate two bodies in contact with each other. Much of the early work in tribology was involved with the force required to separate two metallic surfaces *in static contact*. Bowden and Tabor (1954) at the Physics and Chemistry of Surfaces Laboratory at Cambridge, England, were responsible for the hypothesis (which is still accepted by many tribologists) that the force of *static friction* between two sliding surfaces is, in fact, the force required to break the 'cold welds' which momentarily form between the contacting high spots of the opposing surfaces. Bowden and Tabor's (1954) hypothesis, which will be described in some detail in Section 1.2 of this book, does not necessarily remain valid for materials which form surface films (e.g. oxide films on steels) and there are some formidable difficulties in extending the hypothesis to *dynamic friction*. Essentially, it will be shown that

adhesion probably plays an important role in wear processes which involve only *mechanical* interactions between two contacting surfaces, so that it is these wear processes that should truly be described as 'adhesive wear'. When chemical reactions are also involved, then it seems reasonable to describe those forms of wear according to the type of film formed *after* 'running-in', for example oxidative wear or, more generally, reactive film wear. Undoubtedly, adhesion is involved during 'run-in'. Quite often, investigators have only run their experiments over relatively short times; that is, over part of the initial 'running-in' period, thereby concluding (sometimes incorrectly) that 'adhesive wear' is characteristic of the materials they had been running against each other.

1.1.9 Grooving, scoring, scuffing and spalling

These are just four of the many words often found in the vocabulary of the materials engineer when he/she wishes to describe what has happened to the surfaces that have failed. 'Galling' is another such word. There is no satisfactory standard usage. Typically, 'scuffing' is the term applied to gears that, quite early in their operational career, fail in a fairly catastrophic manner. 'Spalling' relates to a form of failure in which large flat areas of one or both of the surfaces in contact become detached. 'Scoring' and 'grooving' are merely words describing the appearance of one or both of the wearing surfaces in contact. In this book we will only use these words when we speak of the work of other investigators. In all such cases, however, we will also try to establish the basic wear mechanism. The science of tribology has been hindered by these imprecise terms. It is suggested (in Section 1.4) that the results of sliding and/or rolling wear experiments should be discussed in terms of just two mechanisms, namely severe (i.e. mechanical) wear and mild (i.e. chemically reactive) wear. With this simpler classification, it will be much easier for the reader to understand how and why one uses one or more given physical analytical techniques for a given system.

1.1.10 Dry bearings/solid lubricants

Some of the earliest work with solid lubricants involves graphite. Graphite is a typical solid lubricant since:

- (a) it depends on crystallite re-orientation for its effectiveness in lowering the friction between sliding surfaces;
- (b) it involves a layer-like structure which, when oriented in a certain way relative to the sliding interface, provides surfaces of easy-glide;
- (c) it bonds effectively to most metal/metal oxide surfaces.

Both graphite and molybdenite (MoS_2) are often used as solid lubricants, graphite being used in high temperature/normal atmospheric pressure situations, whilst molybdenite is used in normal temperature/high vacuum situations. Dry bearings typically use composite materials, generally a fibre-reinforced plastic, where the fibre is often composed of graphite 'whiskers' (or 'carbon fibres', as they are usually called). For light loads, the dry bearing is often made of a low-friction polymer (such as polytetra-fluoroethylene, PTFE). The wear rate of PTFE is unacceptably high, which is why it cannot be used for highly-loaded tribo-systems without some fibre reinforcement. We shall see, in Chapter 5, that a high

wear rate is often associated with a low friction coefficient between two surfaces treated with solid lubricants. Hence there is always the problem of lubricant replenishment when using solid lubrication.

1.1.11 Bearings

Practically all tribo-systems involve bearings of some sort or another, whether it be a plain bearing, a rolling contact (ball or roller) bearing or an intermittent sliding contact (such as a hinge, lock, circuit breaker, etc.).

The typical plain bearing consists of a shaft rotating inside a journal of slightly larger diameter than the shaft. When stationary, the load of the shaft acts vertically upon the journal, causing wedge-shaped gaps between the two surfaces on either side of the vertical.

Under these conditions, intermetallic contact is prevented by the boundary lubrication properties of the lubricant, that is by the chain-like molecules adsorbed upon the metal surfaces. When rotation occurs, these wedge-shapes become filled with lubricant and actually support the load. This is known as 'hydrodynamic lubrication' and is not now the subject of much research, the basic principles having been laid down by Reynolds at the end of the last century.

The rolling contact bearing, on the other hand, is a much more complicated situation, since it always involves *slip*, as well as pure rolling. Even the enmeshment of gear teeth involves sliding into the region of pure rolling (i.e. where the pitch lines of both gears are tangential to each other), followed by sliding (in the opposite direction) as the gears become disengaged. These contacts are known as 'non-conforming' contacts (in contrast to the conforming contact of the plain journal bearing), and, because of their geometry, the lubricating film within such contacts possesses special properties implied by the term 'elasto-hydrodynamics'. These concepts will be discussed in Section 1.3 rather than in this sub-section, the main purpose of which is to acquaint the newcomer to tribology with some of the terms constantly used in this book.

1.1.12 Tribo-system

We have seen that both the form of lubrication and the type of wear depends very strongly on the geometry of the *system* being subjected to sliding or rolling. Only recently has the systems approach of Czichos and Salomon (1974) begun to be applied to tribology. These authors strongly recommend that we consider all the characteristics and parameters relevant to tribo-testing in terms of (a) operating variables (load, velocity, temperature, duration of test and type of motion); (b) the tribometer test system (the tribo-elements, lubricant and atmosphere); and (c) the tribometric characteristics (the friction force, friction coefficient, noise and vibrations, the temperature, wear rate and contact conditions).

The authors also include surface characteristics as part of characterizing the tribo-elements. The word 'surface characteristics' can, of course, mean everything necessary to characterize a surface, for example the surface topography, its chemical composition, the hardness of the surface and the layers immediately below the surface films normally formed during sliding. Surface characterization of the tribo-elements before, during and after the interaction between moving

surfaces is, of course, the main reason for writing this book. It is the *changes* in surface characteristics caused by the tribometric characteristics that enable us to understand the friction and wear mechanisms that operate in the tribosystem. Czichos and Salomon (1974) do not *explicitly* mention the division of heat at the contacting interface as an important tribometric characteristic. Presumably one could include this under the general heading of 'contact conditions'. We shall see, in Chapter 6, how heat flow analysis can provide independent information about the validity of our hypothetical surface models and hence the validity of our estimates of surface temperatures.

1.1.13 Rheology

This is the study of the flow of materials under stress. In tribology, the flow properties of liquid lubricants are of primary interest. The tribologist wishes to know how a lubricating fluid will behave in the regions of the contacting surface. He is especially interested in the effects of pressure in the contact region. Unfortunately, the effects of temperature upon the properties of the lubricating fluid within the interfacial region has not received as much attention as the effects of pressure. Blok's 'flash temperature' criterion (1937) might be considered to be the pioneering research into the effects of temperature upon the lubrication of moving machinery, in particular gears. In fact, Blok postulates that a plain mineral oil will fail at a certain critical 'flash' temperature, the value of which can be calculated from the experimental conditions. However, his work can tell us nothing about how the lubricating properties of a mineral oil change up to the point at which it fails to function as an effective lubricant.

Essentially there are two ways of defining the property governing the rheology of lubricants, namely the dynamic viscosity (preferred by engineers who have to design bearings) and the kinematic viscosity (preferred by lubricant manufacturers and users). The dynamic viscosity coefficient (η_0) is the constant of proportionality between the shear stress (τ_f) in a liquid and the velocity gradient ($\partial U/\partial y$) between the stationary surface and the moving surface (as shown in Figure 1.1). This can be written as follows:

$$\tau_f = \eta_0 (\partial U / \partial y) \quad (1.1)$$

For a velocity profile similar to the one shown in Figure 1.1, namely a linear increase in velocity parallel to the surface from zero at the stationary surface to U at the moving surface, we can write $(\partial U / \partial y) = U_m / h_{lub}$, where h_{lub} is the distance between the parallel plates. Hence

$$\tau_f = \eta_0 \frac{U_m}{h_{lub}} \quad (1.2)$$

If τ_f is in pascals (i.e. N/m^2), U_m is in m/s and h_{lub} is measured in metres, then the units of dynamic viscosity will be (Ns/m^2). The unit of dynamic viscosity is still taken to be the poise (P), which was originally designated as '1 gram/centimetre-second'. Hence we may write:

$$1(Ns/m^2) = 10 \text{ poise} \quad (1.3)$$

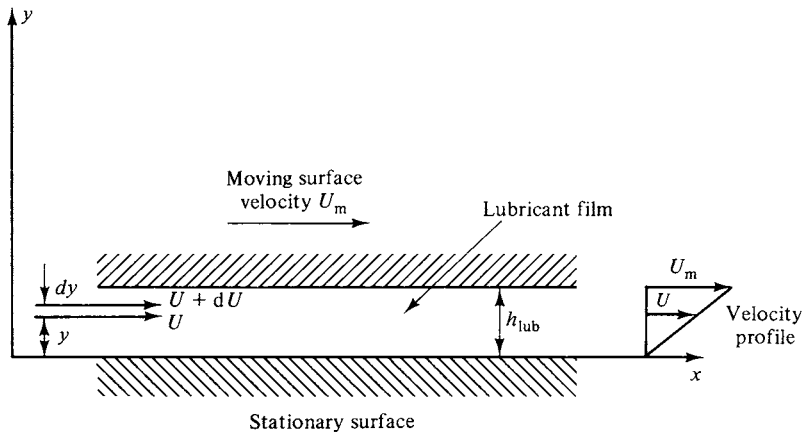


Figure 1.1 *Lubricant film between parallel plates*

In practice, we often talk about ‘centipoise’ (cP) which is (1/100) of 1 P, that is 10^{-3} N s/m^2 . Sometimes, we call the SI unit a ‘pascal-second’, since one (N/m^2) is one pascal. Kinematic viscosity (ν) is defined as:

$$\nu = \eta_0 / \rho \quad (1.4)$$

where ρ is the density of the liquid. We still use the ‘stokes’ as the unit of kinematic viscosity, which was originally designated as ‘ $1 \text{ cm}^2/\text{s}$ ’. Hence we may write the SI unit ($1 \text{ m}^2/\text{s}$) as:

$$1 \text{ m}^2/\text{s} = 10^4 \text{ stokes} \quad (1.5)$$

Again, we tend to use a much smaller unit in practice, namely the ‘centistokes’ (cSt). $1 \text{ cSt} = 1 \text{ mm}^2/\text{s}$ is approximately the value that one obtains for the viscosity of water at 20°C .

Any fluid whose viscosity is independent of the rate of shear (or velocity gradient) is said to be ‘Newtonian’. Equation (1.1) was, in fact, suggested by Newton. Mineral oils and synthetic oils of low molecular weight are Newtonian under almost all practical working conditions. Polymeric liquids of high molecular weight (e.g. silicones, molten plastics) may exhibit non-Newtonian behaviour at relatively low rates of shear. An approximate relation has been proposed for non-Newtonian liquids, namely

$$\tau_f = [\Phi(s_f)]^\beta \quad (1.6)$$

where Φ and β are constants, and s_f is the shear rate. For a Newtonian fluid, $\Phi = 1$ and $\beta = \eta_0$ and $s_f = (\partial U / \partial y)$, so that Equation (1.6) reduces to Equation (1.1) in the limit. For a silicone fluid, the typical value of β is approximately 0.95 (Naylor, 1973).

1.1.14 Summary

In this sub-section, thirteen or so terms have been defined that are in common use amongst practicing tribologists. One of the factors that has hindered progress in tribology must be the imprecise nature of some of the terminology. I have tried to be consistent with my use of these terms within this book. Unless there is a consistent set of definitions, accepted by all concerned, it will become impossible for any technology transfer to occur in this very complex subject. Let us now discuss the basics of the frictional behaviour of materials, since this was, historically, where the first advances were made in tribological knowledge.

1.2 Frictional behaviour of materials

1.2.1 Introduction

The laws of friction have already been mentioned (if only briefly) in this chapter (Section 1.1.2). Let us now discuss these laws more fully. When we try to slide a rectangular block along a plane surface, there is a force of friction between the base of the block and the surface of the plane that opposes motion. If the only other force acting between the block and the plane is the weight of the block (which, in this case, acts at right angles to the general surface between them), then it is found that the external force required to maintain a constant velocity in the plane of motion is always proportional to the normal component of the reaction of the plane to the weight of the block. This external force must, of course, be equal to the kinetic force of friction (F_{kin}). We distinguish between this force and the force opposing that force necessary to start the block moving from the static position, by calling it the static force of friction (F_{stat}). The mechanical engineer, with his/her bias towards rotating machinery and bearings is mainly interested in F_{kin} , whereas it is feasible that the civil engineer would be more interested in F_{stat} .

It is interesting to note that most of the early research has been concerned with F_{stat} . For instance, the pioneer work of the Cambridge School, led by the late Professor Bowden, was almost entirely related to adhesion and the force required to break the cold welds that were said to occur between metal asperities on opposing surfaces in contact. In this section, we will review some of the classical work carried out by the Cambridge School involving plastic deformation and its relation to the laws of friction. This school was mainly responsible for the concept of the real area of contact (A_{real}) being related to the load (W) and the hardness (p_m) of the softer of the two materials in contact, namely:

$$A_{\text{real}} = W/p_m \quad (1.7)$$

We will also review some of the equally classical research of the Aldermaston School, namely the Surface Physics Department of the Research Laboratories of the Associated Electrical Industries group of companies in the United Kingdom. The Aldermaston School was responsible for much of the pioneering research work in what was later called 'tribology'. In particular, their alternative explanation of the laws of friction in terms of purely elastic deformation is well worth being given another 'airing'! This sub-section also describes some of the

later work, carried out in other laboratories, which goes some way towards producing an expression containing both plastic and elastic parameters. Finally, some recent work carried out at the Georgia Institute of Technology in Atlanta is briefly reviewed insofar as it provides experimental evidence that possibly we should *not* be using Equation (1.7) as our basis for the real area of contact while the surfaces are actually in sliding motion.

1.2.2 Plastic deformation of surfaces and its relation to the basic laws of friction

Any hypothesis of friction must take into account the fact that even the smoothest surface obtainable by mechanical polishing comprises many valleys and hills, as revealed by profilometry (see Section 2.3), optical microscopy (Section 2.1) and scanning electron microscopy (see Section 2.3). Hence when two surfaces are in stationary contact over an *apparent* (or nominal) area of contact, they are, in fact, in contact only where the tips of the hills (or 'asperities', as these irregularities are called) on both surfaces actually touch. Clearly, this area must be an area less than the apparent area of contact. We call it the *real* area of contact (A_{real}), already mentioned in Section 1.2.1. Without defining this area more closely for the present, let us consider what the area of contact should be for the idealised case of a single point contact of a hemispherical asperity on a smooth flat surface. Let E_1 and E_2 be the elastic (Young's) moduli for the asperity and flat surface respectively and assume the flat surface is made of softer material than that of the spherical asperity. If now the asperity and the flat are brought into (stationary) contact under a normal load $(W_e)_s$, they will *at first* both deform *elastically*. This means that the spherical cap of asperity material will deform into a flat, whilst the flat will deform into a depression, with both surfaces being in contact over a *circular* region of radius (a_e) given by the following equation:

$$a_e = 1.1 \left(\frac{(W_e)_s(R)}{2} \left(\frac{1}{E_1} + \frac{1}{E_2} \right) \right)^{1/3} \quad (1.8)$$

where R is the radius of the spherical asperity. This equation can be derived from the classical *elastic* deformation equations of Hertz (see, for instance, Holm (1946) or Timoshenko (1934)). From Equation (1.8), we can deduce that the area of contact due to elastic deformation only is proportional to $[(W_e)_s]^{2/3}$, that is

$$(A_e)_s = \pi a_e^2 \propto [(W_e)_s]^{2/3} \quad (1.9)$$

where we have used the subscript 's' to denote that we are considering a *single* contact only and the subscript 'e' to denote an area formed entirely by *elastic* deformation. The mean pressure (p_m) across the single contact is given by:

$$p_m = (W_e)_s / (\pi a_e^2) \quad (1.10)$$

From Equations (1.10) and (1.9), we can readily see that

$$p_m \propto [(W_e)_s]^{1/3} \quad (1.11)$$

This equation tells us that, as we increase the load $(W_e)_s$, the mean pressure increases elastically as the third power of the load. How long can this elastic

deformation continue? The answer is until the softer of the two materials in contact begins to yield. According to Timoshenko (1934), the softer material begins to deform *plastically* (i.e. permanently) at a point P , which is about $(0.6a_e)$ below the surface, when the mean pressure equals $1.1Y$, where Y is the elastic limit. For any *further* increase in load, the deformed region around the initial plastic zone continues to increase in size until it reaches the surface of the softer material (i.e. the flat in this case). The hard indenter then sinks into the softer material until the area of contact is sufficient to support the load. At this stage, Bowden and Tabor (1954) show that the mean pressure (p_m) will be about $3Y$. This is the condition for full plastic flow. Figures 1.2(a) and (b) contain diagrams of (a) the onset of plastic flow and (b) full plastic flow for the single asperity contact model. This is an idealized drawing, but it essentially shows how the area of contact increases elastically up to (πa_e^2) , after which it increases plastically so that p_m becomes identical with mean flow pressure.

We notice that the mean pressure (p_m) is independent of the load (W_p), once full plasticity is obtained. This means that any increase in the load beyond that necessary to obtain (p_m) greater than $3Y$ will only cause the area of contact to increase such that (p_m) remains constant. It is experimentally difficult to verify the predictions relating to the dependence of (p_m) upon load both during elastic and plastic deformation, mainly because plastic deformation generally increases the elastic limit through work-hardening. In fact, it is true to state that most of the basic assumptions made in the plastic deformation theory of friction all tend to be experimentally unproved. Nevertheless, the theory has survived over the last 30 years or so, due mainly to its essential correctness in so far as it tells us how the coefficient of friction depends on the shear strength and hardness of the softer material. Let us briefly list these almost *a priori* assumptions:

- (a) The theory of the *single* spherical asperity on a flat *holds* for the *multiple* asperity contacts of arbitrary geometries.
- (b) The material around contacting asperities is subject to stresses *well beyond* the elastic limit.
- (c) The mean pressure (p_m) for full plastic flow over all the multiple contacts of real surfaces is the load divided by the sum of all the different individual areas of contact between the two surfaces. If we call that sum the real area of plastic, multiple contact $(A_p)_m$, then we have:

$$p_m = \frac{(W_p)_m}{(A_p)_m} \quad (1.12)$$

- (d) When two surfaces come into contact, the harder material sinks into the softer material (of yield pressure p_m) until the area of contact is sufficient to support the load $(W_p)_m$. Under stationary conditions, this entails the 'cold welding' of the mating asperities which are said to be broken when a shearing force equal to the static frictional force (F_{stat}) is applied to the system.
- (e) Welded junctions still form when the surfaces are slid over each other. The force required to make and break the welded asperities during sliding is one component (the adhesive component) of the kinetic friction force (F_{kin}).